LOAN DOCUMENT

PHOTOGRAPH THIS SHEET

DTIC ACCESSION NUMBER

LEVEL

(D)

INVENTORY

NRL-MR-3949

DOCUMENT IDENTIFICATION

15 Mar 79

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

DISTRIBUTION STATEMENT

H A N D L E W I T H C A R E

ACCESSION FOR		
NTIS	GRA&I	☑
DTIC	TRAC	☐
UNANNOUNCED		☐
JUSTIFICATION		
BY		
DISTRIBUTION/		
AVAILABILITY CODES		
DISTRIBUTION	AVAILABILITY AND/OR SPECIAL	
A-1		

DISTRIBUTION STAMP

DATE ACCESSIONED

DATE RETURNED

19960625 190

DATE RECEIVED IN DTIC

REGISTERED OR CERTIFIED NUMBER

PHOTOGRAPH THIS SHEET AND RETURN TO DTIC-FDAC

DTIC FORM JUN 90 70A

DOCUMENT PROCESSING SHEET

PREVIOUS EDITIONS MAY BE USED UNTIL STOCK IS EXHAUSTED.

LOAN DOCUMENT

27 JUN 1996

NRL Memorandum Report 3949

Correlation Algorithms in Naval Ocean Surveillance

H. L. WIENER

Systems Research Branch
Space Systems Division

March 15, 1979

NAVAL RESEARCH LABORATORY
Washington, D.C.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NRL Memorandum Report 3949	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CORRELATION ALGORITHMS IN NAVAL OCEAN SURVEILLANCE		5. TYPE OF REPORT & PERIOD COVERED Interim report on a continuing NRL problem.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) H. L. Wiener		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Research Laboratory Washington, D.C. 20375		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NRL Problem B01-30A 62712N XF12-143-310
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Electronic Systems Command ELEX 350 Washington, D.C. 20360		12. REPORT DATE March 15, 1979
		13. NUMBER OF PAGES 13
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Ocean surveillance	Correlator-trackers
Data processing	Operations research
Algorithms	Evaluation
Correlation	Measures of effectiveness

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Naval Research Laboratory is carrying out a study to collect and evaluate mathematical and computer-based routines for performing the correlation function in Naval ocean surveillance. The correlation function, generally embedded within computerized correlator-trackers, concentrates on determining whether certain elements of the surveillance data base--contact reports or tracks--relate to the same platform. The goal of correlation activities is to partition the data base, so that each set which is a member of the partition is the collection of all data base elements which relate
(Continues)

20. Abstract (Continued)

to a specific platform. The major product of this project is a handbook of correlation schemes, in which such schemes are listed, compared, and evaluated. Specific operations research questions within this project include the determination of measures of effectiveness for comparing correlation routines and the evaluation of routines which have been run on simulated data.

CORRELATION ALGORITHMS IN NAVAL OCEAN SURVEILLANCE

This report constitutes the text of a paper presented at a contributed papers session of the Operations Research Society of America (ORSA) national meeting in November 1978. The paper outlines a Naval Research Laboratory (NRL) project to characterize analytic and computer-based processes for carrying out the correlation function in Naval Ocean Surveillance. Subsequent to presenting the paper the author received a large number of requests for copies. Believing that the paper will be of interest to an even greater audience we are publishing it in the present form.

The NRL Correlation Handbook Project has as its goal the development and publication of a handbook of correlation schemes and algorithms which are applicable to Naval Ocean Surveillance. Because of increased interest in Ocean Surveillance the topic of surveillance data processing and target correlation and tracking has received increased attention recently. References (a) and (b) discuss the multitarget tracking problem and describe current technical developments. This paper discusses the Correlation Handbook project, presents some observations on correlation routines we have encountered, and describes aspects of the correlation problem which lend themselves to operations research analysis.

The correlation handbook project involves characterizing and documenting the current state of the art in surveillance correlation algorithms, evaluating proposed correlation schemes, and identifying needed developments. The Correlation Handbook itself will be an annual document summarizing the results of our investigations. The first edition is due early in 1979.

As used here, the term correlation refers to an activity within surveillance data processing. Other terms which describe ocean surveillance data processing activities are tracking and multisensor interaction. Correlation processes have generally been thought of as schemes for determining when pairs or sets of elements in a data base have a specified relation to each other. For example, it is often desired to ascertain whether two contact reports refer to the same target platform. It may be premature at this time to attempt a precise definition of the term "correlation." We can, however, provide a characterization of correlation processes. In general, these processes

Note: Manuscript submitted January 5, 1979.

(1) Measure the degree of association within and among sets of elements; and

(2) Use specified rules to make inferences or decisions based on the association measurements.

Activities carried out under item (1) frequently involve extensive computations such as Kalman filters and likelihood calculations. Those carried out under item (2) often involve ordering computed association measurements or determining whether computed data exceed some statistically established threshold, so that elements can be associated with each other.

In this project, we are concentrating on examining the correlation function within the context of computerized ocean surveillance data processing systems, of the type called correlator-trackers. These systems operate on data elements such as contact reports or stored track data. We are interested in algorithms developed for all target platforms of Navy interest, undersea, surface, and air. We are, however, initially concentrating on undersea and surface platforms. We cannot investigate correlation by itself because correlation routines are embedded in tracking systems, and it is not always possible to surgically separate correlation modules from other tracker functions. We are devoting a good deal of our attention to correlation schemes which appear in operational or operationally oriented tracking schemes. Thus, in addition to mathematical aspects of the correlation problems we are also paying close attention to the operational aspects.

In carrying out this project we are attempting to meet the needs of at least three different groups, the R&D management community, the development community, and the community of users of correlator-trackers. The first of these needs information to help in planning programs, the second needs information on the state of the art, and the third needs information on what to expect from developmental efforts.

Our efforts comprise five interrelated tasks, collection, analysis, evaluation, comparison, and state-of-the-art assessment. The collection task involves obtaining information on correlation schemes and disseminating information on collected documents. We have a library of more than 60 primary items which we have collected to date. The items in the collection range from individual journal articles to multi-volume reports on computerized processors.

In addition to the library we are preparing a set of one-to-two page descriptions of all relevant documents. These provide brief descriptions of the techniques addressed in the papers, together with indications of the developmental status, important assumptions, or possible difficulties.

Part of the project involves a technical analysis of the field of correlation. The goal is a mathematical characterization of correlation

within the surveillance function. Although the assumptions and mathematical tools vary from situation to situation, we presently feel that the mathematical theory of clustering, in particular partitioning, is the most appropriate model for characterizing correlation activities.

We are engaged in evaluating proposed correlation schemes, to determine their possible worth. The evaluation task is more extensive than that involved in preparing the short summaries. Evaluation of a correlation scheme requires in-depth technical analysis to specify such items as the most important assumptions, the operational situations in which the scheme would be useful, and the potential performance in an operation environment.

Comparison of correlation schemes involves both analysis of technical details and numerical measures of performance against realistic data sets. One current problem is that a complete set of appropriate quantitative measures of effectiveness has not been determined.

State-of-the-art assessments for correlation are necessary not only for developers, but also for R&D managers, who need to be apprised of developmental areas which need emphasis and ones which are relatively mature.

In carrying out this project we have directed our efforts toward developing a structure for the field of correlation as it presently exists, and in doing so identifying general trends and areas of commonality among correlation schemes. A basic requirement is that we develop the proper analytic framework for the project. To expand on an earlier comment, we have determined that an appropriate orientation for our study is to regard correlation as an attempt to partition the set of elements in the data base, whether they be contact reports, tracks, or combinations of these. Each instance of a correlation process represents an effort to conjoin elements, according to some relation which it is desired to represent. The most immediate example of such a relation is that the elements have come from the same target platform. The association of the data base elements according to such a relation is ideally a partition of the set of elements, where each member of the partition contains all elements relating to some platform.

One noticeable trend among correlation-trackers developed for operational contexts is the use of stepwise procedures in carrying out correlations. This type of procedure is based on treating easily analyzed cases initially and then moving through the data correlation problem by considering increasingly difficult cases. For a given set of data, a preliminary pass through the data is used to identify those cases where correlation is immediate, at which time the correlated elements are removed from further consideration. The remaining elements are then analyzed according to different criteria; those elements which are then associated are removed from consideration and the remainder

are again exposed to a different correlation process. The number of sequential steps this procedure can take is determined by the assumed nature of the target set and by the nature of the observed data. The following illustrations constitute a heuristic example for indicating why a stepwise procedure is often valuable. Figure 1 shows a set of target tracks in a crowded environment. At first examination an overall pattern may be difficult to determine. However, it is possible to identify a subset with reasonably regular behavior, for example straight lines, as shown in Figure 2. Removing these from the data set leaves a smaller family for analysis, as shown in Figure 3.

There is a number of assumptions and techniques common to many correlator trackers. One common technique is the use of Kalman filters for updating track data with new observations, and for developing estimates of track position at present and future times. A common assumption which underlies many of the interpretation or decision-making parts of correlation processes is that observations of target location are subject to errors with some form of Normal distribution. Correlation tests are generally of a form which examines the degree of closeness between two elements in the data base. Assumptions of underlying Normal distributions permit the use of the Mahalanobis distance for developing hypothesis tests. It is not clear if these assumptions are based on experimental data, on generally accepted analytic practice, or on computational convenience. To the extent that these assumptions do not affect correlator performance, they need not be investigated too deeply; but should they prove to be important, some degree of validation ought to be done on them.

Up to this time we have described the Naval Correlation Handbook project and some of our observations on correlator-tracker development as it exists today. What we have not yet done is discuss those aspects of the project which have a particular operations research or management science flavor. We propose to do this now.

Two of the major aspects of the project are description and evaluation. We want to be able to describe our findings in terms of most value to the intended audience. The major problem is how to structure the field so that we will be communicating the maximum information. We want to be able to develop families of topics of most value to R&D managers, to algorithm developers, and to algorithm users, and to specify within these families the most useful sets of descriptors for characterizing the schemes. One difficulty is that the intended audiences all have different points of view and therefore different interests. The R&D managers may want to know about the past performance of specific mathematical techniques, such as adaptive Kalman filters or statistical hypothesis testing, whereas algorithm users may be interested in problems of implementation, such as computer-dependence, running time, or the size of the set of platforms with which a correlator can work. The problem of structuring the field, to satisfy the needs of a number of different audiences, is a typical one for OR/MS analysts, and has not yet been completely solved.

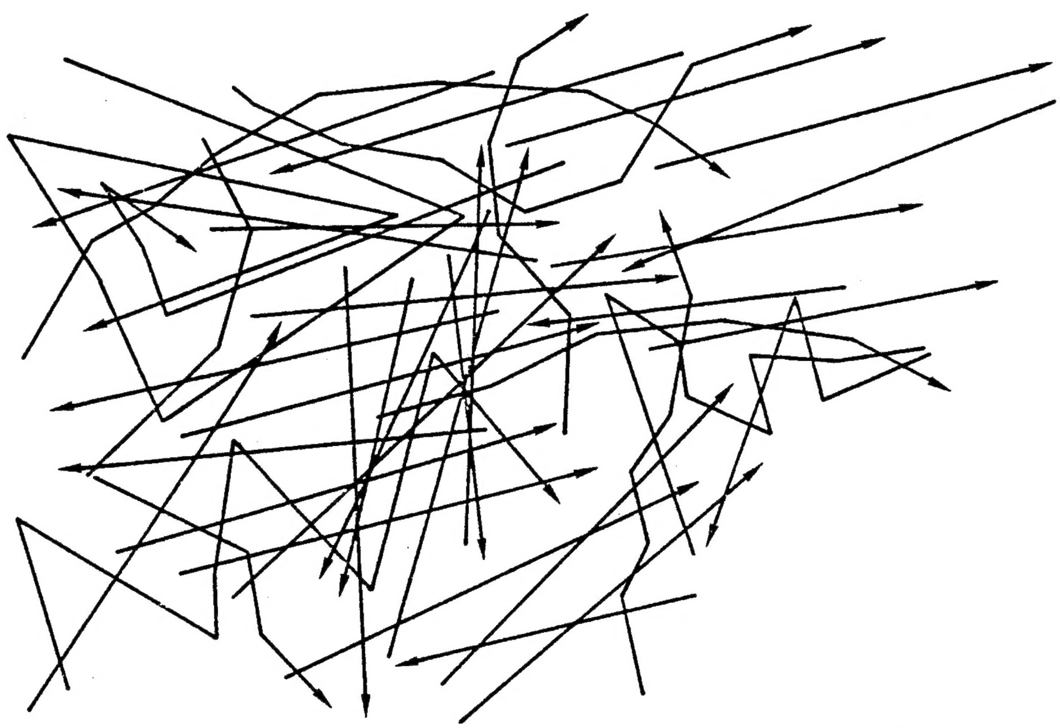

Fig. 1 — Example of ship track histories

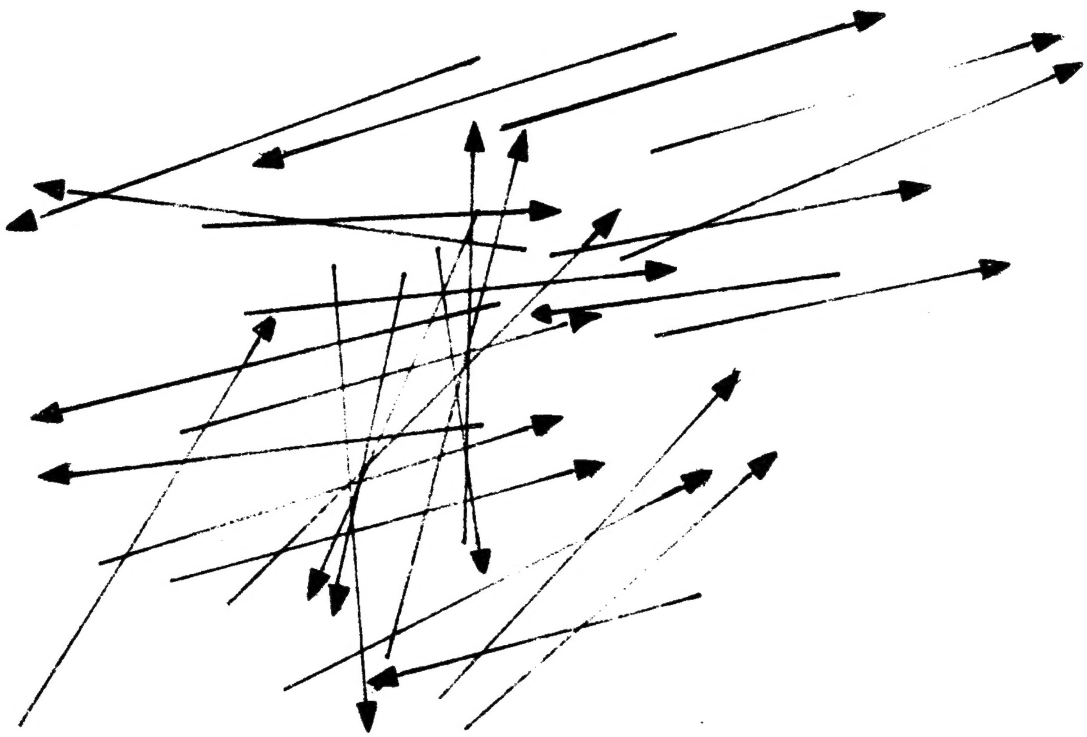

Fig. 2 — Example of ship track histories
"Well Behaved" component

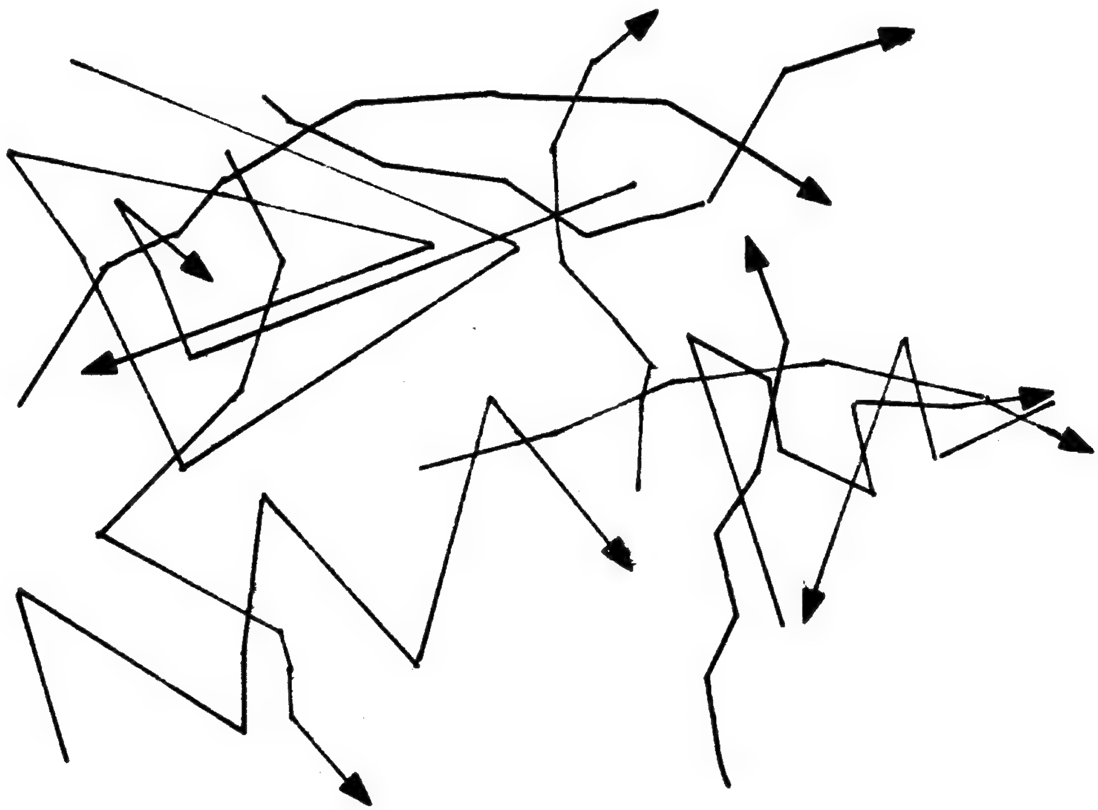

Fig. 3 — Example of ship track histories
"Less Well Behaved" component

The specification of methods for evaluating correlation schemes is an area in which much work still needs to be done. First there is a number of obvious immediate tasks that have to be done, for example, identifying those schemes which are, for all practical purposes, subsumed within other schemes, and identifying those schemes which have serious conceptual errors. The latter task is not done solely by examination of the documentation. It is frequently the case that conceptual errors begin to exhibit their presence when attempts to implement a correlation scheme fail.

A more fundamental problem is the development of measures of effectiveness (MOEs) or "scoring rules" for correlation schemes that are candidates for operational implementation. These measures can be used to evaluate each algorithm separately or can be used to compare algorithms. It is commonly felt that some sort of input data are needed on which to base the evaluation, but there is no general agreement on the best type of data nor on the MOEs whose values can be computed based on algorithm performance on these data. For example, the best test of an algorithm lies in its performance against real-world data, that is, data obtained from operational surveillance sensor systems. A difficulty with these data is that they are not accompanied by corresponding "ground truth" for comparison. The false detection rates and missed detection probabilities are not known, nor is the number of actual platforms. Consequently, it is difficult to conceive of good quantitative measures that truly characterize an algorithm's performance. Some indication of what is happening to those data may be obtained from "symptomatic" measures such as the number of reports which a processor can accept before breaking down, or the proportion of reports which the process judges to be singletons, either false alarms or possible starting points for new tracks. Low values for the first measure or high values for the second may indicate that a processor is not behaving acceptably. These measures are however not diagnostic in the sense that they can be used to determine exactly what is wrong.

Another possibility is to use "canned" data, or simulation-based data which attempt to represent the anticipated input report stream. Such data sets are useful because they permit the development of such numerical scores as the number of correct target tracks established. They also provide a mechanism for adjusting an algorithm's components for enhanced performance. However, it is not always the case that the underlying simulated data really represent the way that actual data would look. Thus, the algorithm may become tuned to an incorrect model. In many instances, the fundamental problem of validating the simulation-based data has not been addressed.

A fundamental problem involving the quality of a correlator-tracker's generated information is how sensitive such information is to the correlation schemes which are used. It is first necessary to develop measures of effectiveness which relate to the use of the information, not just to the computational performance of the algorithms.

It is then necessary to determine how significantly these measures are changed by changing the correlation algorithms. It is generally accepted that a "perfect" correlation scheme will produce "optimum" correlator-tracker information. Analysis of the value of a correlator-tracker must start with an assessment of just how useful such optimum information will be; it should then determine the operational effects which occur as the correlation schemes are changed.

In addition to problems associated with evaluating correlation algorithms, there are also many specific correlation problems which can be addressed from the viewpoint of the mathematics of operations research.

One frequently occurring problem is the correlation of multiple new reports with multiple established tracks, in the sense that certain reports could have come from a number of tracks and certain tracks could have given a rise to a number of reports. A basic form of this situation is illustrated in matrix form in Figure 4, where a measure of association for each report-track pair is assumed to have been given and the goal is a set of unique report-track assignments.

In general, however, there are conceptual problems involved, such as determining both the most appropriate association measures and the criteria for pairing reports with tracks. In Figure 4 the criterion is based on optimizing the sum of the association measures of the report-track pairs and the problem formulation is that of the classical assignment problem. As we mentioned, one commonly used association measure is the Mahalanobis Distance; a correlation criterion which has been employed is to select the set of report-track pairings which minimizes the sum of the distances. Other criteria might be used, for instance selecting those pairings which minimize the probability of no more than two incorrect associations. Once the criteria and association measures have been selected, there are problems involved in developing efficient computational routines such as those which can associate large numbers of candidate tracks and reports in short periods of time.

<u>Summary</u>. In this talk we have described the Correlation Handbook project, indicated certain aspects of correlation within the framework of Naval Ocean Surveillance, and provided examples of open problems which are within the scope of OR/MS techniques. Those who are interested in this area are invited to visit with us and make use of the documents we are collecting. Moreover, and more importantly from our point of view, we would welcome any suggestions you might have on appropriate documentation or on correlation algorithm developers we should contact. Readers desiring further information on the project or its findings should contact the author at Area Code 202, 767-2003.

ESTABLISHED TRACKS

NEW REPORTS

$$
\begin{array}{c|cccc|}
 & 1 & 2 & 3 & 4 \\
\hline
1 & A_{11} & A_{12} & A_{13} & A_{14} \\
2 & A_{21} & A_{22} & A_{23} & A_{24} \\
3 & A_{31} & A_{32} & A_{33} & A_{34} \\
4 & A_{41} & A_{42} & A_{43} & A_{44} \\
\hline
\end{array}
$$

OBJECT: PAIR EACH REPORT I WITH A TRACK J(I) SO THAT:

(A) THE MAP $I \rightarrow J(I)$ IS 1-TO-1.

(B) $\sum_{I} A_{I,J(I)}$ IS OPTIMIZED

Fig. 4 — Multi-report, multi-track correlation problem matrix formulation exhibiting measures of association

References

a. Y. Bar-Shalom, Tracking Methods in a Multitarget Environment, IEEE Transactions on Automatic Control, Vol. AC-23, No. 4, August 1978.

b. C. L. Morefield and C. M. Peterson, "Data Association Algorithms for Large Area Surveillance," paper presented to ORSA/TIMS Meeting, May 1978.

UNCLASSIFIED